National Climatic Data Center

DATA DOCUMENTATION

FOR

DATA SET 6421 (DSI-6421)

Enhanced hourly wind station data for the contiguous United States

December 12, 2002

National Climatic Data Center
151 Patton Ave.
Asheville, NC 28801-5001 USA

:

Table of Contents

Topic	Page Number
1. Abstract	3
2. Element Names and Definitions:	3
3. Start Date	7
4. Stop Date	7
5. Coverage	7
6. How to order data	7
7. Archiving Data Center.	7
8. Technical Contact	7
9. Known Uncorrected Problems	7
10. Quality Statement	9
11. Essential Companion Data Sets	10
12. References	10

1. Abstract: During the NOAA OGP-funded project *"Homogeneous blended wind data over the contiguous United States"* we (NCDC) compiled an hourly wind data set from 1655 stations over the lower 48 states for the period of record (up to year 2000), collected and digitized the station metadata related to the history of the anemometer elevation at most of these stations, and blended the station records with historical snow on the ground information. All this has allowed us to generate homogeneous time series of wind speed at the 10-meter height above the ground/surface for the entire period of digital record (usually, since 1948 for the First Order Stations, since early 1970s for other sites with complete metadata, and since the ASOS implementation at the sites without metadata).

We noted that anemometer elevations throughout the U.S. stations varied widely with time. During the past 60 years, there were up to 12 significant anemometer height changes at some of these stations, and on average there was one change per decade at any station with more than 10 years of record. For example, at Los Angeles International Airport the anemometer height changed 4 times during the past sixty years, varying from 59 ft to 20 ft, while at Edwards Air Force Base, the anemometer height was changed 10 times and varied from 13 to 75 ft. Therefore, the elevation homogenization of the near-surface wind time series is a necessary pre-requisite for any climatological assessments.

In this process we were using the formulae:

$$U_{10g} = U_a \log[(10-H_{snod})/z_0]/\log[(H_a - H_{snod})/z_0], \text{ and}$$
$$U_{10s} = U_a \log[10/z_0]/\log[(H_a - H_{snod})/z_0],$$

where z_0 is the surface roughness (a function of the presence of snow cover at the site); H_{snod} is the snow depth; H_a is the anemometer height above the ground; U_a is the wind speed at the anemometer height, U_{10g} is the speed at 10 meters above the ground, and U_{10s} is the speed at 10 meters above the surface.

The left panel of Figure 1 shows the map of those stations for which we were able to perform the homogenization for the period of record. The right panel of Figure 1 shows the stations for which the homogenization for the pre-ASOS era was impossible due to the absence of metadata (blue dots), and those stations for which any instrument metadata are absent in the NCDC archives. Figure 2 provides general information about the data availability in the archive. Table 1 shows an example of the metadata created for one of the stations.

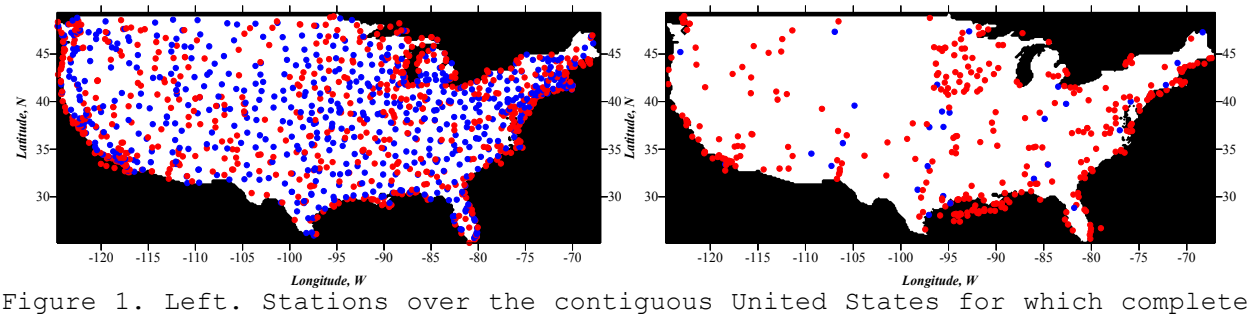

Figure 1. Left. Stations over the contiguous United States for which complete metadata with anemometer information are available (and thus have been digitized) at the National Climatic Data Center, U.S. Air Force Combat Climatology Center, and the U.S. Navy Fleet Numerical Meteorology and Oceanography Detachment in Asheville, North Carolina. Right. Stations over the contiguous U.S. for which

wind information is available in the NCDC data archives in digital form but metadata information with station history is absent. Blue dots indicate the ASOS stations among these stations on both maps.

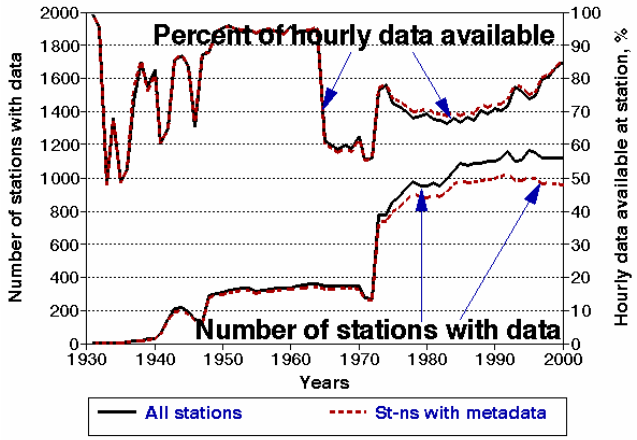

Figure 2. Annual variations of the number of stations with valid data in the archive and the mean percent of valid hourly data available at these stations. A station was considered present in a given year when it has at least 300 valid hourly wind speed values. Stations that do not have at least 1000 valid hourly records during the entire 1931-2000 period were skipped in this graph (90 of them, including 37 stations with metadata).

2. Element Names and Definitions:

The first group of files contains station metadata:

"US_1657stns.lst"
"US_1657stns.lst_expanded"
"metadata_NCDC_stn.lst"
"exist-metadata_files"
"anem_elev_inf".

The original list of stations with hourly information is collected in files **"US_1657stns.lst"** and **"US_1657stns.lst_expanded"**. The format of these two files is similar except:

File **"US_1657stns.lst"** contains supplementary text information in lines after 1662 and in positions after 70 that identify composite stations and provides some additional information on how these stations were handled. The first four lines in **"US_1657stns.lst"** have clarification information about the column meanings, and the last lines (after line 1662) have comments related to different merging steps that were made to several station files during their pre-processing. File **"US_1657stns.lst_expanded"** does not contain this information and has strictly 1657 lines with information (line per station). It (a) has an additional bit of information in the end of each line (blank, or '1') that indicates if the station was in DSI-3280 for at least part of the period ('1' stays for "yes" in this inquiry) (b) shows first and last years with the data for a given station, (c) shows a regional partition of the station within the contiguous U.S. (1 through 9) used in Groisman et al. (2001), and (d) has information about the date of ASOS installation at the station (if any).

Variables included are as follows:

:

An **index** (with values '1 'or '2') that characterizes the number of stations that historically had this ID;

An extended (6-digit) **WMO number** used by USAF for station identification;

The **station name** (up to 19 characters);

A two character **state abbreviation**;

The **longitude** and **latitude** (negative values mean W and S);

The **station elevation** in *meters*;

A 5-digit **U.S. National Weather Bureau number** (WBAN number if present; otherwise blank);

A 6-digit **U.S. cooperative station number** of the station (if present, otherwise blank);

A **time zone indicator** relative to UTC (EST is -5 UTC, PST is -8);

A **one-character indicator** of the first order station (set to '1', otherwise blank; absent in "**US_1657stns.lst**");

The **first and the last year** with valid data in the archive (absent in **"US_1657stns.lst"**);

A **region number** in the following order: Northwest, Missouri River Basin, Upper Mississippi, Northeast, California and Nevada, Southwest, South, Midwest, and Southeast (variable is absent in "**US_1657stns.lst**"; map of this partition is shown in Groisman et al. [2001]);

A **blank** if ASOS was not installed at the station (as of July 31, 2001); otherwise first four characters of this string are set to 'ASOS' and the last 8 characters contain the date (year, month, day) of the ASOS commissioning that can be read in format (i4,2i2.2); this string is absent in "**US_1657stns.lst**".

Example of file **"US_1657stns.lst_expanded"**:

```
1 722025 FT LAUDERDALE/HOLLY FL  -80.15 26.07    7 12849        -5   1972 2000 9 ASOS 19980909
1 722026 HOMESTEAD AFB       FL  -80.38 25.48    2 12826        -5   1943 2000 9
1 722027 MIAMI BEACH (CGS)   FL  -80.13 25.77    1              -5   1983 1984
```

The list of stations with adjusted wind information is provided in "**exist_metadata_files**" which is a subset of the **"US_1657stns.lst_expanded"**.

Variables included are as follows:

The extended (six character) **WMO station number**;

The **station call letters** (if exists, otherwise blank);

The **station coordinates** (lat. and long.);

The **station elevation** (in feet!!!);

The **station name** (up to 35 characters);

The **true WMO #** (if exists; otherwise blank);

A **4-character indication** if the station became ASOS (or AWOS[1]) type;

The **date** of the **ASOS commissioning** (Year, Month, Day);

Up to two WBAN numbers assigned to the station (if any).

Example:

```
722313 LIX  N 30 20 W  89 49    27 SLIDELL, LA         72233       12916
722314 ARA  N 30 02 W  91 53    24 NEW IBERIA, LA    ASOS 19980505 03934
```

File "**anem_elev_inf**" was generated from several supplementary files (available upon request). File "**elev_NCDC_stns**" contains information about the dates of the anemometer elevation changes for all stations with complete metadata digitized at NCDC. The same structure of the anemometer elevation information files was generated for the 260 U.S. first order stations of contiguous U.S. and Alaska [file "**elev_1st_order_stns**"] and for USAF and Navy stations metadata files "**elev_air_force_stns**" and "**elev_navy_stns**". Finally, all this elevation information was combined into the file "**merge_anem_elev_inf.OUT**" in the order of preference (trustworthiness): First are "**elev_1st_order_stns**", then are "**elev_NCDC_stns**", then "**elev_navy_stns**" and "**elev_air_force_stns**". A separate file with duplicate information "**metadata_files_duplicates**" was generated and was used to verify the metadata information about the anemometer elevation that comes from different sources. The discrepancies in this information were analyzed and corrected.

Variables included are as follows:

For each station a line of information contains:

The extended (six-character) **WMO station number**;

The **number of changes** at the station;

Up to 30 groups of changes: **Date of change**: Year, month, day

A **new anemometer elevation** *in feet*.

If the anemometer elevation is not known then it is indicated to be 999.0. If it is known that the AWOS installation was implemented but the anemometer elevation was not reported, then it was assumed to be 10 m (or 33 ft). Complete information about the anemometer elevation at all U.S. ASOS sites was kindly provided by Richard Parry (NOAA/NESDIS). When unknown elevation or day of change were found within the station history record when valid wind speed data exist, the following assumptions were used: (1) If the day of change is

---

[1] If the station became an Automatic Weather Observing station (AWOS, usually around 1990) and later was converted to an ASOS site, then only ASOS type is indicated by a four-character abbreviation "ASOS" or "ASOs".

not known, it is set to the 15th of the month. (2) During the processing of the metadata file for each station, we assumed that the anemometer elevation prior to the first and after the last record of its changes remain the same as indicated in these first and last records respectively. (3) When adjusting wind speed values for the period with unknown anemometer elevation (999.0), it was assumed to be typical for the year when the change to this unknown elevation was documented: 50 feet (year<1951), 20 feet (1959< year <1995), and 30 feet otherwise.

Example of file "**anem_elev_inf**":

```
.....
723125  4 19420915999.0 19510119 30.0 19671115 20.0 19981104 33.0
723126  3 19430715999.0 19440807 39.0 19520415 49.0
723139  8 19460415999.0 19461212 47.0 19550410 56.0 19570301 57.0 19570715 33.0 19601115 32.0 19790415 33.0 19950701 32.8
723140  6 19340810 86.0 19510426 41.0 19540804 58.0 19601117 20.0 19821116 33.0 19980701 32.8
723143  3 19620924 40.0 19720402 35.0 19900111 33.0
723145  7 19480203 30.0 19490528 32.0 19551215 33.0 19700806 22.0 19860411 33.0 19970827 26.0 19970905 32.8
723150  3 19310101 92.0 19640901 20.0 19960601 26.0
.....
```

All files in this group are sorted in ascending order by WMO key.

3. Start Date: 19310101

4. Stop Date: 20001231

5. Coverage: North America

a. Southernmost Latitude: 24N
b. Northernmost Latitude: 50N
c. Westernmost Longitude: 125W
d. Easternmost Longitude: 67W

6. How to Order Data:

Ask NCDC's Climate Services about the cost of obtaining this data set.
Phone: 828-271-4800
FAX: 828-271-4876
E-mail: **NCDC.Orders@noaa.gov**

7. Archiving Data Center:

National Climatic Data Center
Federal Building
151 Patton Avenue
Asheville, NC 28801-5001
Phone: (828) 271-4800.

8. Technical Contact:

National Climatic Data Center
Federal Building
151 Patton Avenue
Asheville, NC 28801-5001
Phone: (828) 271-4800.

9. Known Uncorrected Problems:

We found that the data with several WMO identifiers can be merged together because they are from the same location or from a site that had a small

relocation. The merging has been performed and all composite stations are listed in file "**US_1657stns.lst**" at the last lines (those after 1664). The same file has alternative IDs with the word "merging" in the line with composite stations starting from 70th position. The presence of merged time series should not be considered as a problem ***per se***. But, we sometimes could not trace any logic in partition of data from single station between different WMO station IDs: in some cases originally, different hours of the same day were found in different files.

Several stations with different WMO IDs have the same WBAN numbers, which indicates that they also could be merged. We found that all but three of these stations have data from different periods of time often with a lengthy interval of "no data" between these periods. Only one pair represents a distinctively different neighbor locations. The list of these stations (grouped in pairs with the same WBAN numbers) is below. This list was retrieved from file **"US_1657stns.lst_expanded"**, preserves its structure, and is followed with a short line with commentaries when it is appropriate.

```
1 690070 FORT ORD/FRITZSCHE& CA -121.77 36.68   41 93217 043186  -8   1960 1993 5
1 724916 FORT ORD/FRITZSCHE  CA -121.77 36.68   41 93217 043186  -8   1973 1991 5
OVERLAP. But, sometimes there are slightly different values

1 722223 PENSACOLA REGIONAL& FL  -87.18 30.47   37 13899         -6   1990 2000 9 ASOS 19971201
1 722220 PENSACOLA           FL  -87.20 30.47   36 13899 086997  -6 1 1948 2000 9 ASOS 19971201
OVERLAP. But, sometimes there are slightly different values

1 722635 BORGER              TX  -99.75 30.50  522 13973 414671  -6   1972 1977 7 ASOS 19961202
1 747400 JUNCTION (AMOS)     TX  -99.77 30.50  522 13973         -6   1977 2000 7 ASOS 19961202

1 722746 SAFFORD MUNICIPAL   AZ -109.63 32.85  968 93084         -7   1972 1973 6 ASOS 19970903
1 722747 SAFFORD (AMOS)      AZ -109.68 32.82  950 93084         -7   1978 2000 6 ASOS 19970903

1 723139 HATTERAS (ASOS)     NC  -75.62 35.22    3 93729         -5   1996 2000 9 ASOS 19950701
1 723040 CAPE HATTERAS       NC  -75.55 35.27    3 93729 311458  -5 1 1957 2000 9 ASOS 19950701
OVERLAP.  But, sometimes slightly different values

1 723193 WINSTON-SALEM/SMITH NC  -80.22 36.13  296 93807         -5   1990 2000 9 ASOS 19981202
1 723190 WINSTON-SALEM/SMITH NC  -80.23 36.13  298 93807 319539  -5   1972 1985 9 ASOS 19981202

1 724243 LONDON-CORBIN ARPT& KY  -84.08 37.08  369 03849         -5   1991 2000 8 ASOS 19960918
1 723290 LONDON              KY  -84.07 37.08  369 03849 154898  -5   1972 1985 8 ASOS 19960918

1 724360 COLUMBUS/BAKALAR    IN  -85.88 39.25  201 13803         -5   1984 1984 8
1 724363 COLUMBUS/BALKALAR & IN  -85.90 39.27  200 13803         -5   1990 2000 8

1 724370 TERRE HAUTE/HULMAN  IN  -87.30 39.45  181 93823         -5   1972 1985 8
1 724373 TERRE HAUTE/HULMAN& IN  -87.32 39.45  178 93823         -5   1990 2000 8

1 725118 HARRISBURG/CAPITAL& PA  -76.85 40.22  106 14751         -5   1948 2000 4 ASOS 20001011
1 725110 HARRISBURG/CAPITAL  PA  -76.85 40.22  106 14751 363699  -5   1972 1985 4 ASOS 20001011
OVERLAP.  Values are very close and mostly supplement each other

1 725896 LAKEVIEW            OR -120.35 42.22 1455 24270         -8   1977 1997 1
1 725891 LAKEVIEW            OR -120.35 42.18 1455 24270         -8   1976 1977 1

1 726459 STURGEON BAY      & WI  -87.32 44.78  176 04824         -6   1981 1996 3
1 726458 STURGEON BAY        WI  -87.42 44.85  221 04824         -6   1981 2000 3
These are two different stations with distinctively different metadata.

1 726460 WAUSAU/ALEXANDER    WI  -89.62 44.92  367 14897 478968  -6   1972 1985 3 ASOS 20000914
1 726463 WAUSAU MUNICIPAL  & WI  -89.63 44.93  366 14897         -6   1990 2000 3 ASOS 20000914

1 726954 NEWPORT             OR -124.07 44.58   48 24272         -8   1990 2000 1
1 726951 NEWPORT             OR -124.05 44.63   48 24272         -8   1976 1977 1

1 727927 OCEAN SHORES (CGS)  WA -124.13 46.97    3 94225         -8   1972 1990 1 ASOS 20010322
:
:
```

```
1 727923 HOQUIAM/BOWERMAN    WA -123.93 46.97    4 94225 453807  -8   1991 2000 1 ASOS 20010322

1 744200 ROOSEVELT         & UT -109.98 40.30 1556 24084 427395  -7   1985 1997 6
1 725708 ROOSEVELT           UT -109.98 40.30 1556 24084         -7   1972 1985 6

1 745200 MILFORD             UT -113.02 38.42 1533 23176         -7   1983 1985 6 ASOS 19960801
1 724797 MILFORD             UT -113.00 38.35 1508 23176 425654  -7   1997 2000 6 ASOS 19960801
```

The data set has more than 7000 hourly wind speed values that are flagged as grossly erroneous. The user is strongly recommended not to neglect these flags.

ASOS issues. An important homogeneity issue, which has not yet been resolved, was an assessment of the changes in extreme wind speeds with transition to ASOS. This issue is additionally complicated with an algorithm change (so-called firmware 3.0 introduction). This firmware was gradually introduced since 1995. Both ASOS implementation and firmware replacement directly affect the averaging period used in the hourly wind speed evaluation and, while being irrelevant for mean wind speeds (McKee et al. 2000), changes extreme wind speed estimates. We found that the frequencies of calm weather and extreme wind speed were significantly affected after the ASOS implementation. We also found that the ASOS stations prior to the firmware change (the exact date of which is not always well known) behave quite differently compared to the post-firmware change period as well as with the same period but at the non-ASOS stations. Thus we recommend excluding these data from any possible change assessments that involve extreme and low wind analyses.

Prior to winter 1948-1949, the snow on the ground information from U.S. cooperative stations is sparse everywhere except the Midwestern U.S. and, therefore, we often had to use long-term snow climatology in our formulae (described in Abstract) instead of actual daily snow depth information to derive the 10-meter wind speed (variables **Sp10** and **Sp_surf**).

When the day of the anemometer height change was not known (but month and year were shown), it was assumed to be the 15th of this month. Furthermore, in each day of the anemometer height change at each station, we assumed that the change has occurred at 00 LST and used the new height in adjustments of the wind speed to 10 m above the ground/surface.

10. Quality Statement: This data set has undergone extensive quality checks on all parameters, including range checks for wind direction and extreme winds assessment. Questionable data have been flagged.

For each state we were originally determined to find at least 10 highest wind speed hourly values that could be used as a benchmark for further quality control of these and forthcoming data as well as for practical use. For this purpose for each state, we selected the ten maximum wind speed values (at the anemometer height) and carefully analyzed them by comparing with the entire suite of weather information during the entire day at the neighboring stations including the wind extreme location. Trajectories of Atlantic hurricanes were also used during this analysis. When we saw that the extreme value is unbelievably high compared to those during the neighboring hours, at the neighboring sites, and under the weather conditions in the area (atmospheric pressure and temperature gradients, cloudiness, weather codes, and presence of hurricane/tropical storm/depression nearby were considered), we flagged this value as "questionable" (flag "Q") and continued our search for extremes. Quite often during this analysis, we found that the value most probably has a wrong scale (i.e., multiplied by 10) and scaling back will return it to a proper value similar to those in the neighborhood. We flagged these values as

"with scale-error" (flag "S") and also continued our search for extremes. During this analysis we also looked for pattern of errors and tried to reveal them within the entire station (group of stations) records. It appears that no one among the first "extremes" was a true value but a result of some or another type of error. We intended to identify at least 490 extreme wind speed values (10 for each state within the contiguous U.S. but 20 for New Hampshire, where Mt. Washington station represents a group of its own with extremely high winds throughout the entire period of observation) and assign them one of two flags: (1) valid extreme well supported by the neighbor information "E" and (2) valid extreme somewhat supported by the neighbor and weather information "U". But, prior to identification, we found approximately 7,000 individual wind speed values that were flagged as "Q" or "questionable" and/or "S" or "with scale-error". Furthermore, we checked all wind speed values that were above 30 m sec^{-1} throughout this process because we believe that whatever the maximum is, the speeds above 30 m sec^{-1} are too damaging and should be also checked. We also had to exclude 16 pieces of hourly information (from week to up to three years of data) from the station data that have multiple errors of unknown origin (e.g., every second day some of hourly records were replaced with very high wind speed, let say, 30 m/sec, from a given direction (always the same) that is accompanied with negative temperatures (even in summer). These types of errors were discovered mostly on unmanned stations. About eighty wrong extreme wind speed values were discovered in the data originated from DSI-3280, while the bulk of others came from the DATSAV3 data stream. This was expected because of the highest standards of quality assessment used in the DSI-3280 data stream. However, it also should be noted that DSI-3280 represent only ~15% of the stations (also they provide the longest time series) in our data collection. Quality flags are an integral part of each individual wind speed value in the archive.

11. Essential Companion Datasets: None. But, acquaintance with documentation in DSI-3280 and DSI-3505 would be helpful.

12. References:

Court, A. 1953: Wind extremes as design factors. *J. Franklin. Inst.*, 256, No.1, 39-55.

DeGaetano, A.T., 1998: Identification and implications of biases in U.S. wind observation, archival and summarization methods, *Theoretical and Applied Climatology*, 60, 151-162.

Golubev, V.S., and Bogdanova, E.G. 1996: Accounting of blowing snow events in precipitation measurement in Russia. *Proc. of the ACSYS Solid Precipitation Climatology Project Workshop*, Reston, VA, USA, 12-15 September 1995, World Meteorol. Organ., WMO/TD No. 739.

Groisman, P.Ya., D.R. Easterling, R.G. Quayle, V.S. Golubev, A.N. Krenke, and A.Yu. Mikhailov, 1996: Reducing biases in estimates of precipitation over the United States: phase three adjustments. *J. Geophys. Res.*, 101, 7185-7195.

Groisman, P.Ya., Easterling, D. R., Quayle, R. G., and Golubev, V.S., 1998: Response to the Comments by D. Yang and B.E. Goodison on the paper by P.Y. Groisman et al. "Reducing biases in estimates of precipitation over the United States: Phase 3 adjustments". *J. Geophys. Res.*,103, 6229-6235.

Groisman, P.Ya., R.W. Knight , and T.R. Karl, 2001: Heavy precipitation and

high streamflow in the contiguous United States: Trends in the 20th century. *Bull. Amer. Meteorol. Soc.*, 82, 219-246.

Lockhart, T.J. 2000: Summary of wind climate continuity with ASOS. Proc. of the AMS 12th Conf. On Appl. Climatology, 8-11 May, 2000, Asheville, NC, 131-133.

McKee, T.B., N.J. Doesken, Ch. A. Davey, and R.A. Pielke, Sr., 2000: Climate Data Continuity with ASOS. Report for period April 1996 through June 2000. Colorado Climate Center, Fort Collins, Colorado, 25 pp.

Quayle, R.G., and R.G. Steadman, 1998: The Steadman wind chill: An improvement over present scales. *Weather and Forecasting*, 13, 1187-1193.